The PAR promoter expression system: modified lac promoters for controlled recombinant protein production in Escherichia coli

Hothersall, Joanne; Godfrey, Rita; Fanitsios, Christos; Overton, Tim; Busby, Steve; Browning, Doug

DOI:
10.1016/j.nbt.2021.05.001

POSTPRINT

The PAR promoter expression system: modified *lac* promoters for controlled recombinant protein production in *Escherichia coli*

Joanne Hothersall [1]*, Rita E. Godfrey [1], Christos Fanitsios [2,3], Tim W. Overton [2], Stephen J. W. Busby [1]*, and Douglas F. Browning [1,4]*

[1] Institute of Microbiology and Infection and School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.

[2] School of Chemical Engineering and Institute of Microbiology and Infection, University of Birmingham, Birmingham, B15 2TT, UK.

[3] Present address: Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.

[4] College of Health and Life Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK.

*To whom correspondence should be addressed:

Email: j.hothersall@bham.ac.uk Tel: +44 (0)121 414 5434

Email: s.j.w.busby@bham.ac.uk Tel: +44 (0)121 414 5439

Email: d.f.browning@bham.ac.uk Tel: +44 (0)121 414 5434

Highlights.

- **The PAR promoter system can express recombinant proteins to many different levels.**
- **The system is versatile, functioning in many *E. coli* strains and growth regimes.**
- **Promoters are tightly regulated, allowing low-level expression of toxic proteins.**

ABSTRACT

Many commonly used bacterial promoters employed for recombinant protein production (RPP) in *Escherichia coli* are capable of high-level protein expression. However, such promoter systems are often too strong, being ill suited for expressing proteins that are difficult to fold or proteins that are targeted to the membrane or secreted out of the cytoplasm. To circumvent this we have constructed a suite of bacterial promoters with a range of different promoter strengths, assigning them specific promoter activity ratings (PARs). Selecting three of these PAR promoters, with low, intermediate and high strengths, we demonstrate that the expression of target proteins, such as green fluorescent protein (GFP), human growth hormone (hGH) and single chain variable region antibody fragments (scFvs) can be set to three levels when expressed in *E. coli*. We show the PAR promoter system is extremely flexible, operating in a variety of *E. coli* strains and under various different culture regimes. Furthermore, due to its tight regulation, we show that this system can also express a toxic outer membrane protein, at levels, which do not affect bacterial growth. Thus, the PAR promoter system can be used to tailor the expression levels of target proteins in *E. coli* and maximize RPP.

Abbreviations: GFP, green fluorescent protein; hGH, human growth hormone; HRP, horseradish peroxidase; IPTG, isopropyl β-D-1-thiogalactopyranoside; LacI, *lac* operon repressor; MCS, multiple cloning site; PAR, promoter activity rating; RPP, recombinant protein production; scFv, single chain variable region antibody fragment.

Keywords: Recombinant protein production, *Escherichia coli*, transcription regulation, *lac* promoter, membrane proteins

Introduction

The bacterium *Escherichia coli* has been the cornerstone of recombinant protein production (RPP) for many years. *E. coli* is particularly well suited to its role as a protein production factory as it grows quickly to high density in cheap medium and can be **easily manipulated genetically** [1]. As a consequence, many recombinant protein expression systems have been constructed, such as those based on the *E. coli* lactose operon promoter (*e.g.* the *lac* and *tac* promoters) and the T7 RNA polymerase expression system [2, 3, 4]. In such systems, the gene of interest is usually cloned into a plasmid vector downstream of a strong regulated promoter and RPP is initiated by the addition of an inducer molecule, *e.g.* the lactose analogue, IPTG (isopropyl β-D-1-thiogalactopyranoside) [5]. In many instances, these systems allow the production of large amounts of high quality protein. However, as RPP expression systems typically have strong inducible promoters, problems can arise when expressing difficult-to-fold targets, membrane proteins and proteins secreted out of the cell or into periplasm. Thus, high-level expression of such proteins often leads to product misfolding, resulting in target degradation or its aggregation into inclusion bodies [6, 7, 8, 9]. In an attempt to increase the amount of soluble target, RPP is often carried out at lower temperatures, or expression levels are decreased by using a weaker promoter or lower inducer concentrations [6, 10, 11]. This often involves trial and error, and may even require switching the expression system, since low-level RPP with some systems, *e.g.* the T7 RNA polymerase expression system, can be hard to control.

Ideally, RPP expression systems should be as versatile as possible, being tightly regulated, allowing high- and low-level target expression, and be compatible with many different *E. coli* strains and "helper" plasmids, which can carry various tRNAs or chaperones [1, 10]. Above all, expression systems should be easy to use, with vectors carrying multiple cloning sites (MCS), purification and secretion tags, and, potentially, different antibiotic resistance cassettes. Previously we generated a suite of promoters, which were based on the *lac* and *tac* promoters, and assigned them promoter activity ratings (PAR values) based on their strengths **[5]**. These PAR promoters (PAR1 to PAR8) show a wide range of promoter activities, stretching from low-level (*i.e.* PAR1) to high-level (*i.e.* PAR8) expression capabilities [5]. Each PAR promoter also carried *lac* operator sequences and were, therefore, repressed by the *lac* operon repressor (LacI) and IPTG-inducible (Fig. 1a). As some of these promoters appeared to be promising for RPP, we have transferred them to easy-to-use vector backbones to generate the PAR promoter expression system. By picking three promoters, which have low, intermediate and high expression capabilities, when fully induced by IPTG, this system is able to express recombinant proteins at three different levels in *E. coli*, allowing the expression level of a target protein to be easily tailored to maximize the production of soluble recombinant

protein. In addition, we demonstrate that the PAR promoter system can be tuned by different IPTG concentrations, can be used to express toxic proteins, and is flexible, functioning in different *E. coli* strains, media and at different growth temperatures.

Materials and methods

Bacterial strains, plasmids, and materials

E. coli strains, XL1 Blue and JCB387, were used for plasmid construction and DNA manipulation, whilst *E. coli* BL21, BL21(DE3), W3110 and SHuffle Express and were used for recombinant protein over-expression (Supplementary Table S1). Strains were grown in LB broth (Sigma), Lennox broth (2% (w/v) peptone (Oxoid), 1% (w/v) yeast extract (Oxoid) and 170 mM NaCl) [12] and auto-induction medium [13], with appropriate antibiotic selection (ampicillin 100 µg/ml, kanamycin 50 µg/ml). For RPP, *E. coli* BL21, BL21(DE3) and W3110 were routinely grown at 37°C, whilst SHuffle Express was grown at 30°C.

Expression vector construction

The PAR promoters, PAR1 to PAR8 (including PAR4L, formerly *lac*O3O1) have been described previously (Browning *et al*., 2019). Each PAR promoter was amplified from plasmids pRW50/ PAR1 to PAR8 (Supplementary Table S1), using PCR with primers detailed in Supplementary Table S2. Purified PCR products were restricted with BglII and XbaI and cloned into the pET22b and pET26b expression vectors (Novagen), replacing the canonical T7 RNA polymerase promoter (Supplementary Fig. S1). The DNA encoding 6His-GFP, from pET15b/ *6his-gfp*, was cloned into each pET22b and pET26b PAR construct using XbaI and BamHI restriction sites (Supplementary Table S1; Supplementary Figs. S2 and S3). The DNA encoding hGH-6His and anti-IL-1β-6His scFv, from pHAK1 and pYU49, respectively, was cloned into each pET22b PAR construct using NdeI and SacI (Supplementary Figs. S2 and S3) [14, 15]. The DNA encoding the $BamA_{ENm}$ chimeric outer membrane protein was cloned into pET22b PAR1, using NdeI and XhoI (Supplementary Figs. S2 and S3) [16]. The *lacI*^{q} mutation, which changes a single base in the *lacI* promoter to increase its strength [17, 18], was introduced into pET22b PAR7 and pET26b PAR7 derivatives using the Agilent QuikChange site-directed mutagenesis kit and primers lacIqF/R (Supplementary Table 2). All constructs were verified by Sanger DNA sequencing.

Recombinant protein over-expression and detection

Bacterial cultures of *E. coli*, carrying pET expression plasmids containing the PAR promoters and various target genes, were grown with shaking in 10 mL of LB medium, until an optical density (OD_{600}) of 0.3 to 0.5. Protein over-expression was induced by the addition

of IPTG and samples were taken after three or four hours induction. For the expression of the $BamA_{ENm}$ chimera, BL21 cells were grown in 50 ml Lennox broth, supplemented with glucose, where indicated. Total protein samples were routinely prepared by resuspending normalized amounts of cells in 2X Laemmli loading buffer (Sigma), heating at 95°C for three minutes, and centrifuging prior to loading. Normalized protein samples were resolved by reducing SDS-PAGE and analyzed using Coomassie blue staining and Western blotting, as in our previous work [19]. For Western blotting, 6His-GFP was detected using anti-GFP antiserum raised in mouse (Sigma) and an anti-mouse-HRP secondary antibody (Sigma), hGH-6His was detected using anti-hGH antiserum raised in rabbit [5] and an anti-rabbit-HRP (horseradish peroxidase) secondary antibody (Amersham), and anti-IL-1β-6His scFv was detected using anti-6His (C-terminal)-HRP (Invitrogen). $BamA_{ENm}$ was detected using anti-BamA antiserum [16] and an anti-rabbit-HRP secondary antibody (Amersham). Blots were developed using Pierce ECL Western blotting substrate and all gels and blots shown are representative. To access the aggregation of product in inclusion bodies, total, soluble and insoluble protein samples were also prepared using Agilent BugBuster, according to the manufacturer's instructions.

Cellular fractionation and membrane isolation

50 ml cultures of BL21 cells, carrying pET22b PAR1/ $BamA_{ENm}$, were grown in Lennox broth, supplemented with 0.2% glucose, in the presence or absence of 1 mM IPTG for three hours. Cells were isolated by centrifugation and pellets were washed with 10 mM Tris-HCl (pH 7.4) and resuspended in 20ml of 10 mM Tris-HCl (pH 7.4) containing 2 mM phenylmethylsulfonyl fluoride (PMSF, protease inhibitor). Cell envelopes were disrupted by continuous passage through an Emulsiflex C3 for 5 minutes, and unbroken cells and particulate material was removed by centrifugation for 15 min at 6,000 x g and 4°C. The total membrane fraction (inner and outer membranes) was then isolated by centrifuging the supernatant for 1 hour at 48,000 x g at 4 °C and the soluble fraction, which contains cytoplasmic and periplasmic proteins was retained [16, 19]. Membranes were washed once and resuspended in 0.4 ml of 10 mM Tris-HCl (pH 7.4).

Rescue of BamA depletion in E. coli

To determine the ability of the $BamA_{ENm}$ chimera to rescue BamA depletion on solid media, the *E. coli* BamA depletion strain JWD3 (Supplementary Table S1) [20] was grown on LB agar plates, supplemented by 0.2% (w/v) arabinose or 1 mM IPTG, where indicated. For experiments in liquid media, JWD3 cells were grown in 50 ml of Lennox broth with 0.2% glucose, with shaking at 37°C, in the presence or absence of 0.05% (w/v) arabinose. Optical density was monitored and after 2 hours of growth (OD_{600} = 0.3 - 0.4) $BamA_{ENm}$ production was induced by the addition of IPTG. The preparation of normalised total cellular protein samples,

after 6 hours of growth, was as in Browning *et al*. **[19]**. Note that in JWD3 cells chromosomally encoded *E. coli* BamA is only produced in the presence of arabinose, whilst in its absence, BamA expression is prevented, resulting in the cessation of growth and cell death [20]. Depletion can be rescued by providing a functional copy of *bamA*, such as that carried by pET22b PAR1/ $BamA_{ENm}$ [16, 19].

Flow cytometry.

For flow cytometry analysis, 50 mL cultures of LB medium were incubated with shaking at 37°C until the culture reached OD_{600} ~0.6, and then RPP was induced by addition of IPTG for three to four hours, as stated. Cultures were analysed using a BD Accuri C6 flow cytometer (Becton Dickinson, UK). Samples were mixed with 0.2 µm-filtered PBS and data was collected at a rate of 1000 - 4000 events per second using slow flow and a forward scatter height (FSC-H) threshold of 10000 to eliminate non-cellular material until 20000 events had been recorded per sample. Data were analysed using CFlow software (BD).

Results and Discussion

Construction and RPP with the PAR promoters

Previously we generated a suite of PAR promoters, which covered a broad range of promoter activities from low to high expression levels (*i.e.* PAR1 to PAR8) (Fig. 1a) [5]. For ease of use, each PAR promoter was sub-cloned into the medium copy number expression vector pET22b, which carries an extensive MCS to facilitate gene cloning and a functional *lacI* gene to ensure regulation in any *E. coli* host strain (Supplementary Fig. S1 and Table S1). To investigate expression from these new vectors, DNA encoding N-terminally 6His-tagged GFP (*6his-gfp)* (Supplementary Figs. S2 and S3) was cloned downstream of each PAR promoter and recombinant plasmids were transferred into *E. coli* BL21 cells. Cells were grown in LB medium until mid-logarithmic growth (OD_{600} between 0.3 and 0.5) and recombinant PAR promoter expression was induced with 1 mM IPTG for three hours. Levels of total GFP production were then analysed by SDS-PAGE and Western blotting. Results in Fig. 1b show that IPTG-induced GFP production increased with the strength of the PAR promoter (*i.e.* from PAR1 to PAR8). Most PAR promoters were tightly regulated, with little or no expression in the absence of IPTG, but some, *e.g.* PAR6 and PAR8, were found to be leaky (Fig. 1).

As we wished to develop vectors with a weak, intermediate and strong promoter, we chose the PAR1, PAR4L and PAR7 constructs, respectively. As expression from the PAR7 construct was slightly leaky in the absence of inducer (Figs. 1b and 1c), the $lacI^q$ mutation, which increases the expression of LacI, was introduced [17, 18]. This new construct, referred to as pET22b PAR7Q, showed minimal 6His-GFP expression in the absence of IPTG, as

judged by Western blotting (results not shown). Expression of 6His-GFP in BL21, driven by the PAR1, PAR4L and PAR7Q promoters, was produced at low, intermediate and high levels respectively, after induction with 1 mM IPTG (Fig. 2a). This was confirmed by flow cytometry (Fig. 2b), which revealed differences in expression after 1 hour of induction.

Since RPP expression systems should be as flexible as possible, the PAR1, PAR4L and PAR7Q promoters were also introduced into the pET26b expression vector, which carries an alternative plasmid backbone and a kanamycin resistant cassette (Supplementary Table S1). Expression studies, again using *6his-gfp*, demonstrated that graded levels of 6His-GFP expression, as expected, were achieved (Supplementary Fig. S4). Furthermore, our three plasmid PAR promoter system functioned well in the *E. coli* K-12 strain W3110 (Supplementary Fig. S5), and with auto-induction medium when cells were grown at different temperatures (*i.e.* 30 and 37°C) for longer periods of time (*i.e.* 23 hrs) (Supplementary Fig. S6) [13]. Thus, we conclude that the PAR promoter system is versatile and can be used to express recombinant proteins to set levels, in different strains and under different growth conditions.

Maximizing the solubility of recombinant 6His-GFP using the PAR promoters

In many instances, high level RPP can result in misfolded proteins and aggregation of product into inclusion bodies [6]. We, therefore, hypothesised that expressing target proteins, using the weaker PAR1 and PAR4L promoters, may reduce product aggregation and improve overall protein solubility. To examine this, we analysed soluble and insoluble fractions from BL21 cells expressing 6His-GFP from the pET22b PAR1, PAR4L and PAR7Q constructs. Results illustrated in Fig. 2c, show that, for the strong PAR7Q construct, a large proportion of 6His-GFP is found in the insoluble fraction. For the intermediate strength PAR4L promoter, less insoluble 6His-GFP is observed, whilst for the weak PAR1 promoter, all the 6His-GFP was found in the soluble fraction. This shows that the different expression levels achieved with the PAR promoter system can be used to tailor expression levels and minimize insoluble product formation. Note, analysis of the quantity of 6His-GFP in the soluble and insoluble fractions correlates with flow cytometry data (Fig. 2b), whereby the fluorescence of the PAR7Q cultures is only slightly higher than the PAR4L cultures despite containing more total GFP. Flow cytometry has been shown to measure both quantity and folding quality of GFP, with insoluble GFP having low fluorescence [11].

Fine-tuning of the PAR promoter response by altering the inducer concentration

In our previous experiments, we used a saturating concentration of IPTG (*i.e.* 1 mM) to ensure that all promoters were fully derepressed. However, for many induction regimes, low

inducer concentrations are used to try to decrease RPP and fine-tune expression levels [6]. Whilst this works for some expression systems, in other systems this leads to only a proportion of the cells in a culture expressing recombinant protein, which has been termed as an all-or-none phenomenon [21, 22, 23]. Therefore, we examined whether 6His-GFP expression from our pET22b PAR1, PAR4L and PAR7Q constructs was tuneable. Once more, BL21 cells, carrying each plasmid, were grown in LB and induced with different IPTG concentrations (*i.e.* 2, 10, 50 and 1000 μM). The expression of 6His-GFP was then monitored using flow cytometry. Results in Fig. 3 show that for all three promoters different levels of expression could be set in a culture by using different IPTG concentrations. Furthermore, the analysis of individual cells indicated that for each promoter and IPTG concentration tested, GFP induction was homogenous within the bacterial cell population (Supplementary Fig. S7). This is particularly evident for the PAR7Q construct, which produces discrete GFP-expressing populations at many different IPTG concentrations, indicating that expression from this highly active promoter can be effectively tuned by different IPTG concentrations.

Expression of different protein targets using the PAR promoter system

To test the versatility of the PAR promoter expression system, we examined the expression of two additional targets, human growth hormone (hGH) and a single chain variable region antibody fragment against interleukin 1β (anti-IL-1β scFv). Thus, the DNA encoding each protein, carrying a C-terminal 6His tag, was cloned into pET22b, carrying either the PAR1, PAR4L or PAR7Q constructs (Supplementary Figs. S2 and S3) [14, 15]. The resulting plasmids were then transferred to BL21 cells and RPP was induced by the addition of 1 mM IPTG to mid-logarithmic growing cells. Results in Fig. 4 show that, as anticipated, graded levels of expression were achieved for both hGH-6His and anti-IL-1β-6His scFv, with the most product produced by cells carrying the PAR7Q construct and the least for PAR1.

As correct folding of hGH requires the formation of a disulphide bond, we examined whether expressing hGH-6His with the PAR promoters aided its solubility. However, as the *E. coli* cytoplasm is a reducing environment that does not favour disulphide bond formation, it was unsurprising to find that the majority of hGH-6His was insoluble (Fig. 5a). To circumvent this problem, hGH-6His expression was carried out in *E. coli* SHuffle Express, a genetically modified *E. coli* strain, which enables cytoplasmic disulphide bond formation. Cells were grown in LB medium at 30°C and RPP induced with 1 mM IPTG for three hours. Results illustrated in Fig. 5b, demonstrate that hGH-6His was successfully induced under this altered induction regime and that for all PAR promoter constructs the majority of recombinant hGH-6His was now found in the soluble fraction. Note that the PAR7Q construct produced the most insoluble

product and that the intermediate strength promoter PAR4L gave the best yield of soluble protein with minimal insoluble protein, as detected by Western blotting (Fig. 5b).

The PAR1 promoter can be used to express toxic proteins

For some target proteins, very low levels of expression are required, especially when the recombinant product is toxic. As our PAR1 promoter is based on the *lac* promoter, it is subject to catabolite repression and can be inhibited by the inclusion of glucose in the growth medium (Supplementary Fig. S8) [24]. Therefore, to test if the PAR1 promoter could be used to express a toxic protein, we cloned the DNA encoding $BamA_{ENm}$, a large chimeric outer membrane protein (OMP) from *Neissieria meningitidis*, into our pET22b PAR1 vector (Supplementary Figs. S2 and S3) [16]. This 88 kDa membrane protein has potential as a vaccine candidate against *N. meningitidis* [16, 25]. Results in Fig. 6a, show that overexpression of this construct at 37°C, using the PAR1 promoter with 1 mM IPTG is toxic and leads to the cessation of cell growth. When a lower IPTG concentration was used (*i.e.* 20 µM) cells reached a higher optical density but, growth was arrested before the end of the experiment (Fig. 6a). In contrast, induction of $BamA_{ENm}$ expression with 1 mM IPTG in the presence of glucose did not influence bacterial growth (Fig. 6a) and resulted in lower expression levels of $BamA_{ENm}$ without toxicity (Fig. 6b; lane 6). Fractionation of cells into their soluble (cytoplasmic and periplasmic proteins) and membrane components (inner and outer membranes) confirmed that that $BamA_{ENm}$ was located in the membrane fraction (Fig. 6c), as expected for an integral outer membrane protein.

In *E. coli*, BamA is an essential protein that is responsible for inserting bacterial β-barrel containing OMPs into the bacterial outer membrane [26]. Previously, we demonstrated that very low-level expression of the *N. meningitidis* $BamA_{ENm}$ chimera could function in *E. coli*, rescuing the depletion of BamA in the *E. coli* K-12 strain JWD3, where BamA production is absolutely dependent upon arabinose [16]. Results in Supplementary Fig. S9 demonstrate that IPTG induced $BamA_{ENm}$ expression from pET22b PAR1, in the presence of glucose, could also rescue depletion of BamA in JWD3, indicating that under these expression conditions, $BamA_{ENm}$ was folded and functional. It is also of note that, in the absence of IPTG, BamA depletion in JWD3 was not rescued and $BamA_{ENm}$ expression was not detected (Supplementary Fig. S9), indicating that the PAR1 promoter is tightly regulated and suitable for the expression of toxic proteins.

Conclusions

Many expression plasmids carry strong promoters. In most instances, this is beneficial, but, in some cases, high-level RPP can result in the accumulation of insoluble protein into inclusion bodies, or cell death if the expressed product is toxic. As obtaining the correct expression levels for problematic proteins can be difficult to achieve, we have developed the PAR promoter system, which consists of three plasmids with low, intermediate and high expression capabilities (PAR1, PAR4L and PAR7Q). Thus, by cloning target DNA into each vector, the most suitable level of expression required for optimal RPP and solubility can be determined quickly. The pET22b and pET26b vectors that we used carry extensive MCS, tags for purification and secretion, and different antibiotic resistance cassettes (Supplementary Fig. S1 and Table S1). Each plasmid also carries the gene encoding the Lac repressor (i.e. *lacI*) and, thus, coordinated regulation can achieved in any *E. coli* strain regardless of its genotype. Consistent with this, we show that the PAR system functions with different *E. coli* strains commonly used in industry and academia (*e.g.* BL21, W3110 and SHuffle Express).

Using the PAR system we have expressed different proteins (GFP, hGH, an scFv and $BamA_{ENm}$) ranging in size from 23 to 88 kDa (Supplementary Fig. S3). Interestingly, even with GFP, which is often used as a model protein for expression analysis, high-level expression with the PAR7Q construct resulted in substantial product insolubility, with the weaker PAR promoters producing less insoluble product (Fig. 2c). Note that expression levels from our PAR7Q promoter construct rivals that of the highly active T7 expression system and so this effect is to be expected (Supplementary Fig. S10). Cytoplasmic expression of more complex proteins, such as hGH, can be more problematic as it requires disulphide bond formation for correct folding. In this instance, product solubility was greatly improved by employing *E. coli* SHuffle Express as an expression host (Fig. 5), as it allows cytoplasmic disulphide bond formation to occur. Once more, the largest amount of insoluble product was found when using our strongest promoter construct, PAR7Q, with little or no product insolubility observed for the PAR4L and PAR1 constructs. Thus, we show, as others before us, that reducing RPP expression, by using weaker promoters can improve target solubility [6, 10].

For many experiments, we used high concentrations of IPTG (*i.e.* 1 mM) to ensure that our expression systems are fully switched on. However, our results show that the level of RPP driven by the PAR promoters can be modulated. Our systems are tuneable, with specific IPTG concentrations producing different expression levels homogenously within a culture, rather than an all-or-none phenotype that has been observed before (Fig. 3 and Supplementary Fig. S7) [21, 22, 23]. Also, the use of glucose-mediated repression with the PAR1 promoter enabled the expression of the toxic $BamA_{ENm}$ chimera from *N. meningitidis* (Fig. 6). It is of note

that $BamA_{ENm}$ is a large outer membrane protein that must traverse the *E. coli* inner membrane and periplasmic space to be inserted into the outer membrane [26]. Overloading of the cellular machinery responsible for these events (*i.e.* the Sec translocase, the periplasmic chaperones, and the nascent Bam complex) will likely result in toxicity and cell death [16, 26]. Thus, due to the tight repression of PAR1 in the absence of inducer and the ability to modulate expression by glucose (Fig. 6; Supplementary Figs. S8 and S9), the PAR1 promoter is ideal for low-level expression of toxic proteins. Finally, all three promoters worked well with auto-induction medium (Supplementary Fig. S6), indicating that inducer exclusion (*i.e.* the ability to prevent lactose uptake when glucose is present in the growth medium) is a feasible way to control and delay RPP induction with the PAR system [13].

Since its discovery, the *lac* operon promoter and its derivatives have been extensively used in biotechnology [3, 4, 8, 27]. In this work, we have further adapted the *lac* promoter to generate an easy-to-use RPP expression system that allows the expression of target proteins to be quickly set to obtain optimal expression and/ or solubility. Furthermore, we show that the PAR system functions well with many of the common induction regimes used to control both the level and timing of target protein expression. Thus, fine-tuning expression levels from the PAR promoters gives added flexibility. Future research will focus on optimizing the PAR promoter system for use in larger-scale expression and fermenter applications.

Acknowledgements

J.H. was generously supported by an Industrial Biotechnology Catalyst (Innovate UK, BBSRC, EPSRC) (BB/M018261/1) to support the translation, development and commercialisation of innovative Industrial Biotechnology processes and by a BBSRC IAA Follow-on-Fund award (BBSRC IAA BB/S506709/1). D.F.B was supported by BBSRC grants BB/M018261/1 and BB/R017689/1. C.F. was supported by a studentship from the BBSRC Midlands Integrative Biosciences Training Programme.

Author contribution

J.H., T.W.O., S.J.W.B. and D.F.B. devised the research programme, experiments were performed by J.H., R.E.G., C.F. and D.F.B., and the manuscript was written by J.H., S.J.W.B. and D.F.B., with input from all authors.

Conflict of interest statement

The Authors declare no conflict of interest.

References

[1] Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 2014;5:172. doi: 10.3389/fmicb.2014.00172.

[2] Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 1986;189(1):113-30.

[3] Amann E, Brosius J, Ptashne M. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene 1983;25(2-3):167-78.

[4] de Boer HA, Comstock LJ, Vasser M. The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc Natl Acad Sci U S A 1983;80(1):21-5.

[5] Browning DF, Godfrey RE, Richards KL, et al. Exploitation of the *Escherichia coli lac* operon promoter for controlled recombinant protein production. Biochem Soc Trans 2019;47(2):755-763. doi: 10.1042/bst20190059.

[6] Overton TW. Recombinant protein production in bacterial hosts. Drug Discov Today 2014;19(5):590-601. doi: 10.1016/j.drudis.2013.11.008.

[7] Wagner S, Klepsch MM, Schlegel S, et al. Tuning *Escherichia coli* for membrane protein overexpression. Proc Natl Acad Sci U S A 2008;105(38):14371-6. doi: 10.1073/pnas.0804090105.

[8] Browning DF, Richards KL, Peswani AR, et al. *Escherichia coli* 'TatExpress' strains super-secrete human growth hormone into the bacterial periplasm by the Tat pathway. Biotechnol Bioeng 2017;114(12):2828-2836. doi: 10.1002/bit.26434.

[9] Quick M, Wright EM. Employing *Escherichia coli* to functionally express, purify, and characterize a human transporter. Proc Natl Acad Sci U S A 2002;99(13):8597-601. doi: 10.1073/pnas.132266599.

[10] Gopal GJ, Kumar A. Strategies for the production of recombinant protein in *Escherichia coli*. Protein J 2013;32(6):419-25. doi: 10.1007/s10930-013-9502-5.

[11] Vera A, González-Montalbán N, Arís A, et al. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol Bioeng 2007;96(6):1101-6. doi: 10.1002/bit.21218.

[12] Squire DJ, Xu M, Cole JA, et al. Competition between NarL-dependent activation and Fis-dependent repression controls expression from the *Escherichia coli yeaR* and *ogt* promoters. Biochem J 2009;420(2):249-57. doi: BJ20090183 [pii] 10.1042/BJ20090183.

[13] Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 2005;41(1):207-34. doi: 10.1016/j.pep.2005.01.016.

[14] Alanen HI, Walker KL, Lourdes Velez Suberbie M, et al. Efficient export of human growth hormone, interferon alpha2b and antibody fragments to the periplasm by the *Escherichia coli* Tat pathway in the absence of prior disulfide bond formation. Biochim Biophys Acta 2015;1853(3):756-63. doi: 10.1016/j.bbamcr.2014.12.027.

[15] Matos CF, Robinson C, Alanen HI, et al. Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in *Escherichia coli* CyDisCo strains. Biotechnol Prog 2014;30(2):281-90. doi: 10.1002/btpr.1858.

[16] Browning DF, Bavro VN, Mason JL, et al. Cross-species chimeras reveal BamA POTRA and beta-barrel domains must be fine-tuned for efficient OMP insertion. Mol Microbiol 2015;97(4):646-59. doi: 10.1111/mmi.13052.

[17] Calos MP. DNA sequence for a low-level promoter of the *lac* repressor gene and an 'up' promoter mutation. Nature 1978;274(5673):762-5.

[18] Muller-Hill B, Crapo L, Gilbert W. Mutants that make more *lac* repressor. Proc Natl Acad Sci U S A 1968;59(4):1259-64.

[19] Browning DF, Matthews SA, Rossiter AE, et al. Mutational and topological analysis of the *Escherichia coli* BamA protein. PLoS One 2013;8(12):e84512. doi: 10.1371/journal.pone.0084512.

[20] Lehr U, Schütz M, Oberhettinger P, et al. C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. Mol Microbiol 2010;78(4):932-46. doi: 10.1111/j.1365-2958.2010.07377.x.
[21] Novick A, Weiner M. Enzyme Induction as an all-or none phenomenon. Proc Natl Acad Sci U S A 1957;43(7):553-66.
[22] Maloney PC, Rotman B. Distribution of suboptimally induces -D-galactosidase in *Escherichia coli*. The enzyme content of individual cells. J Mol Biol 1973;73(1):77-91.
[23] Khlebnikov A, Risa O, Skaug T, et al. Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. J Bacteriol 2000;182(24):7029-34.
[24] Kaur J, Kumar A, Kaur J. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. Int J Biol Macromol 2018;106:803-822. doi: 10.1016/j.ijbiomac.2017.08.080.
[25] Guan Q, Wang X, Wang X, et al. *In silico* analysis and recombinant expression of BamA protein as a universal vaccine against *Escherichia coli* in mice. Appl Microbiol Biotechnol 2016;100(11):5089-98. doi: 10.1007/s00253-016-7467-y.
[26] Konovalova A, Kahne DE, Silhavy TJ. Outer Membrane Biogenesis. Annu Rev Microbiol 2017;71:539-556. doi: 10.1146/annurev-micro-090816-093754.
[27] Makoff AJ, Oxer MD. High level heterologous expression in *E. coli* using mutant forms of the *lac* promoter. Nucleic Acids Res 1991;19(9):2417-21.
[28] Wilson CJ, Zhan H, Swint-Kruse L, et al. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. Cell Mol Life Sci 2007;64(1):3-16. doi: 10.1007/s00018-006-6296-z.

Figure Legends

Fig. 1. Expression of recombinant 6His-GFP protein can be set to different levels using the PAR promoters. (a) The panel shows a schematic representation of the PAR promoters used in this study. Each PAR construct (*i.e.* PAR1 to PAR8) carries -10 and -35 promoter elements, based on either the *lac* or *tac* promoters, **with** flanking *lac* operator DNA sequences, **which are either the O1 or O3 operator sequences** [5]. **Base substitutions are defined by the position of the base altered, with respect to the *lac* transcription start site (+1), and the substituted base introduced.** The LacI repressor protein, binding to each operator target, represses promoter activity (-ve), until the addition of IPTG causes it to release the promoter DNA [28]. Panels (b) and (c) show Coomassie blue stained SDS-PAGE gels and Western blots, respectively, examining 6His-GFP expression in *E. coli* BL21 cells, carrying various pET22b PAR constructs (PAR1 to PAR8). Cells were grown in LB medium and sampled after three hours induction with (or without) 1 mM IPTG. An empty pET22b vector control (EV) was included. In the Western blot in panel (c), recombinant 6His-GFP was detected using anti-GFP antiserum and anti-mouse HRP secondary antibody.

Fig. 2. Analysis of 6His-GFP expression using the three promoter PAR system. (a) The panel shows a Coomassie blue stained SDS-PAGE gel of *E. coli* BL21 cells expressing 6His-GFP, using the pET22b three promoter PAR system (PAR1, PAR4L and PAR7Q). Cells were grown in LB medium and RPP was initiated for three hours by the addition of 1 mM IPTG, where indicated. (b) The panel shows the mean cellular fluorescence of *E. coli* BL21 cells as measured by flow cytometry, expressing 6His-GFP from the pET22b three promoter PAR system. Cells were grown in LB medium and RPP was induced for three hours using 1 mM IPTG. Data are shown as mean green fluorescence values from replica flasks and error bars are ± the standard deviation. (c) The panel shows a Coomassie blue stained SDS-PAGE gel investigating the solubility of 6His-GFP expressed in *E. coli* BL21 cells using the pET22b three promoter PAR system. Cultures were grown in LB medium and protein production was induced by 1 mM IPTG for three hours. Harvested cells were lysed to prepare total (T), soluble (S) and insoluble (I) protein samples. In panels (a) and (b) empty vector controls (EV) were included.

Fig. 3. Expression from the PAR promoters can be fine-tuned using different IPTG concentrations. The figure shows flow cytometry analysis of mean green fluorescence from BL21 cells carrying pET22b PAR/ 6His-GFP constructs (a) PAR1, (b) PAR4L and (c) PAR7Q. Cells were grown in LB medium and 6His-GFP expression was induced by the inclusion of IPTG at 2, 10, 50 and 1000 µM. Data are shown as mean green fluorescence values from replica flasks, error bars are ± the standard deviation.

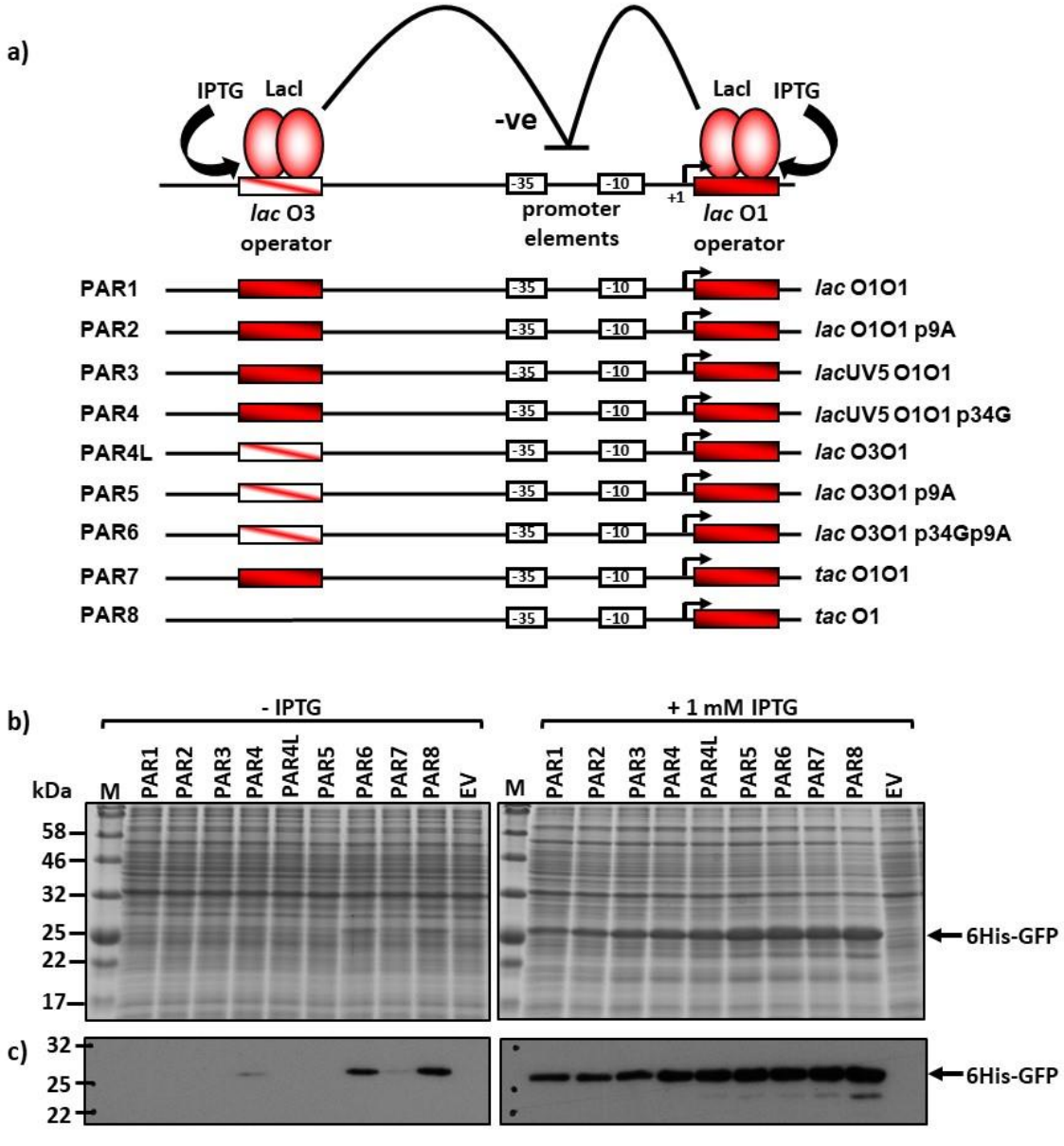
Fig. 4. Expression of hGH and an anti-IL-1β scFv using the PAR promoter system. The figure shows Coomassie blue stained SDS-PAGE gels of *E. coli* BL21 cells expressing (a) hGH-6His and (b) anti-IL-1β-6His scFv, using the pET22b three promoter PAR system (PAR1, PAR4L and PAR7Q). Cells were grown in LB medium and RPP was initiated for three hours by the addition of 1 mM IPTG, where indicated. In each case, an empty vector control (EV) was included.

Fig. 5. Solubility of recombinant hGH-6His expressed in *E. coli* BL21 and *E. coli* SHuffle Express cells. The figure shows Coomassie blue stained SDS-PAGE gels investigating the solubility of hGH-6His expressed in (a) *E. coli* BL21 and (b) *E. coli* SHuffle Express cells using the pET22b three promoter PAR system (PAR1, PAR4L and PAR7Q). Cultures were grown in LB medium and protein production was induced by 1 mM IPTG for three hours. Harvested cells were lysed to prepare total (T), soluble (S) and insoluble (I) protein samples. Empty vector controls (EV) were included. In panel (b) a Western blot is included detailing the detection of hGH-6His in samples, using anti-hGH antiserum and anti-rabbit HRP secondary antibody.

Fig. 6. Expression of the *N. meningitidis* $BamA_{ENm}$ chimera protein in *E. coli* BL21 cells. (a) The panel shows growth of the *E. coli* BL21 cells carrying either pET22b PAR1 empty vector or pET22b PAR1/ $BamA_{ENm}$ in Lennox broth, supplemented with 0.2% glucose (G 0.2%), where indicated. Expression of $BamA_{ENm}$ was induced after three hours growth by the addition of IPTG at either 20 µM or 1 mM, where indicated. (b) Detection of $BamA_{ENm}$ chimera expression. The panel shows a Coomassie blue stained gel of normalised total cell protein from the BL21 cells in panel (a), carrying either pET22b PAR1 empty vector or pET22b PAR1/ $BamA_{ENm}$ after three hours induction. Culture numbering in panel (a) is the same for the loading of samples in panel (b) (denoted as *). The gel was loaded as follows: lane 1, BL21 pET22b PAR1 empty vector (EV) grown in Lennox broth; lane 2, BL21 pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth (uninduced); lane 3, pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth with 0.2% glucose (uninduced); lane 4, BL21 pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth and induced with 1 mM IPTG; lane 5, BL21 pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth and induced with 20 µM IPTG; lane 6, BL21 pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth with 0.2% glucose and induced with 1 mM IPTG. (c) Detection of $BamA_{ENm}$ in membrane fractions from BL21 pET22b PAR1/ $BamA_{ENm}$ cells. The panel shows a Coomassie blue stained gel and Western blot of soluble (Sol) and membrane (Mem) fractions from the BL21 cells in panel (a), carrying pET22b PAR1/ $BamA_{ENm}$ grown in Lennox broth with 0.2% glucose in the presence or absence of 1 mM IPTG after 3 hours. For the Coomassie blue

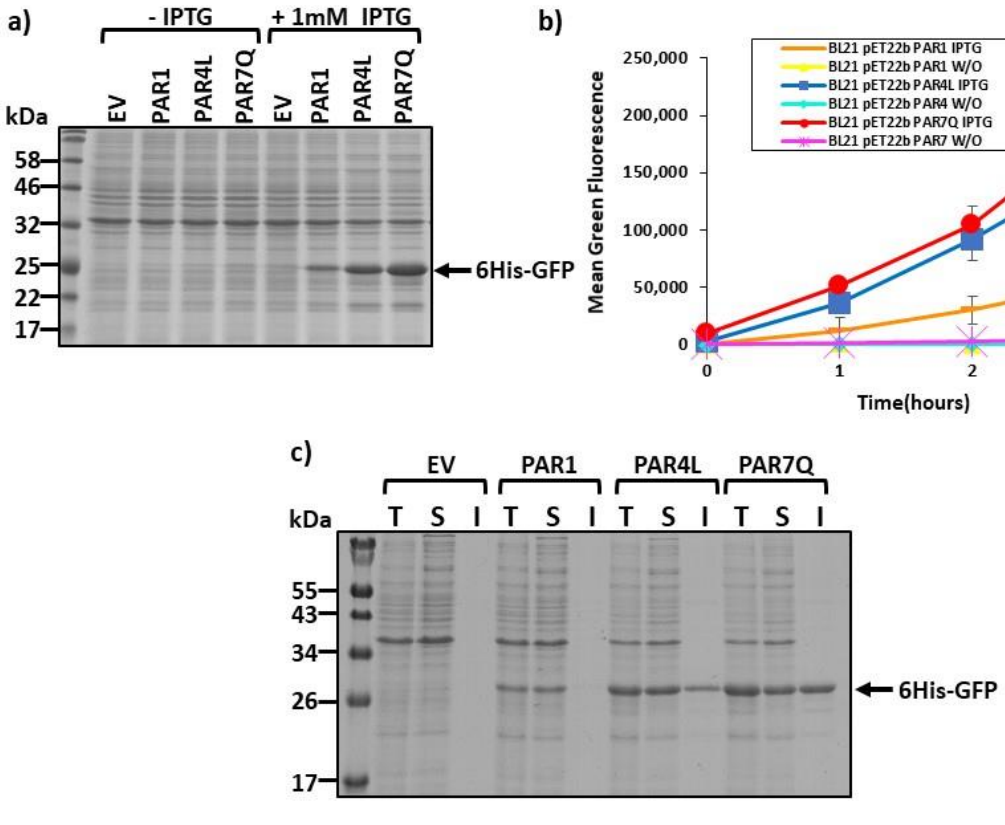
stained gel, 5 µg of soluble and 3 µg of membrane protein was loaded, and for the Western blot 0.5 µg and 0.3 µg of protein were loaded, respectively. $BamA_{ENm}$ was detected by probing with anti-*E. coli* BamA POTRA antiserum and anti-rabbit HRP secondary antibody.

Fig. 1. Hothersall *et al.* (2021)



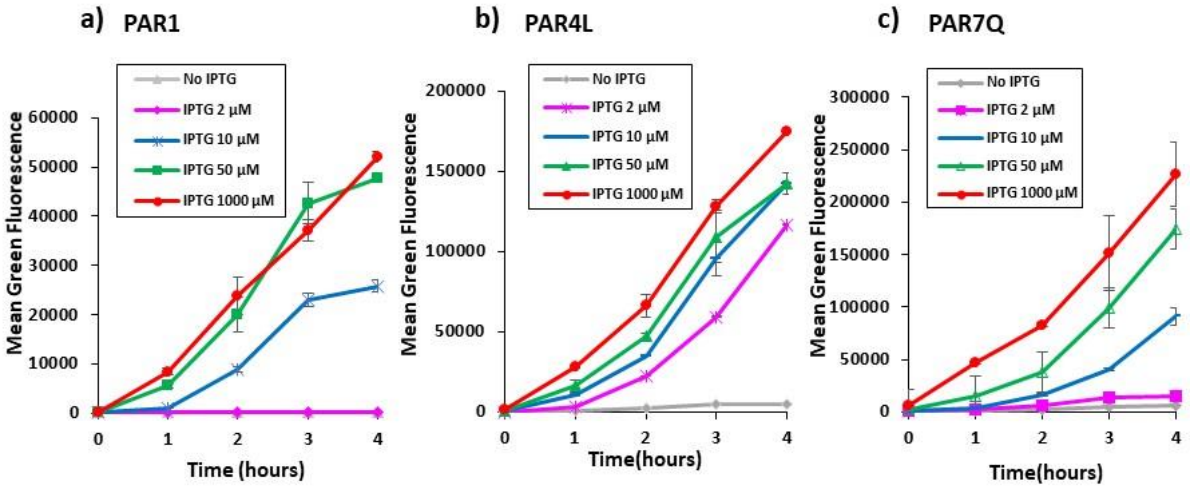
552

Fig. 2. Hothersall *et al.* (2021)



553

Fig. 3. Hothersall *et al.* (2021)



554

555

Fig. 4. Hothersall *et al.* (2021)

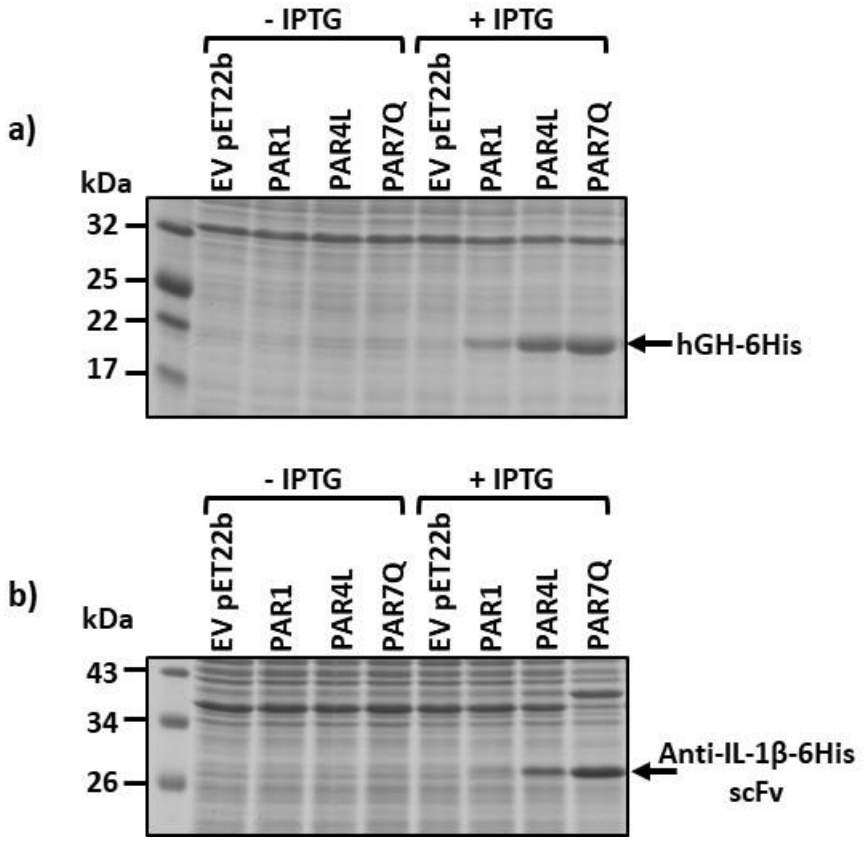
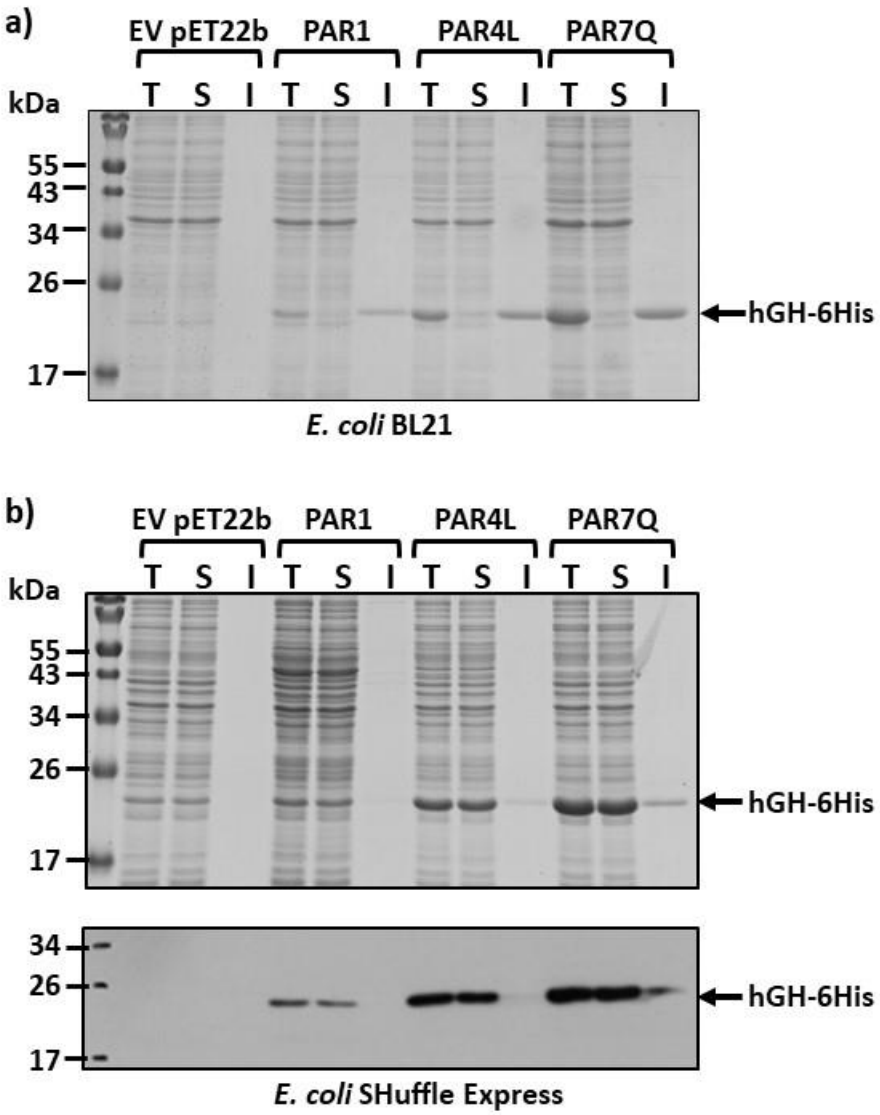
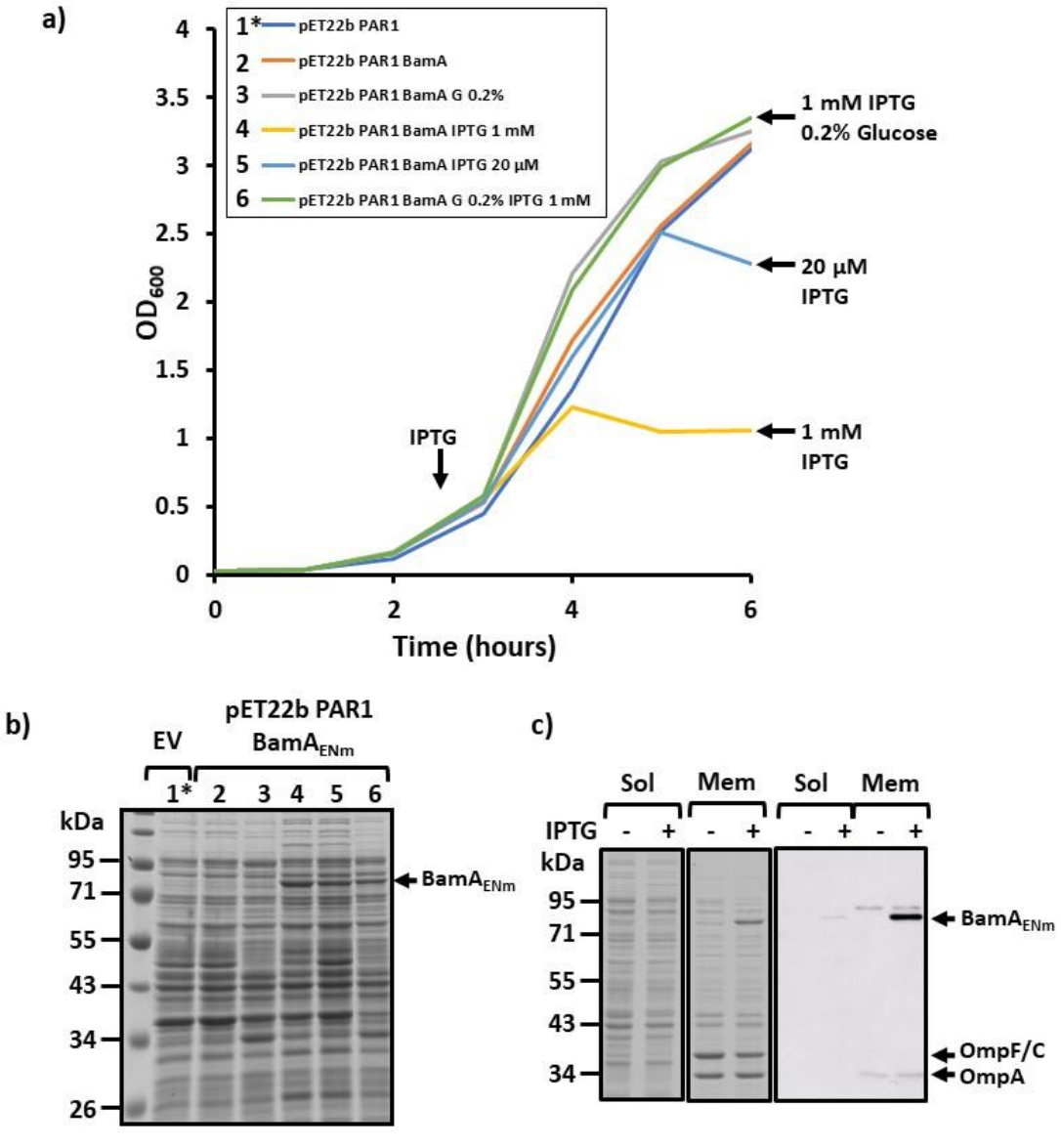


Fig. 5. Hothersall *et al.* (2021)



558

Fig. 6. Hothersall *et al.* (2021)



559

The PAR promoter expression system: modified *lac* promoters for controlled recombinant protein production in *Escherichia coli*

Supplementary Material.

Joanne Hothersall [1]*, Rita E. Godfrey [1], Christos Fanitsios [2,3], Tim W. Overton [2], Stephen J. W. Busby [1]*, and Douglas F. Browning [1,4]*

[1] Institute of Microbiology and Infection and School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.

[2] School of Chemical Engineering and Institute of Microbiology and Infection, University of Birmingham, Birmingham, B15 2TT, UK.

[3] Present address: Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.

[4] College of Health and Life Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK.

*To whom correspondence should be addressed:

Email: j.hothersall@bham.ac.uk Tel: +44 (0)121 414 5434

Email: s.j.w.busby@bham.ac.uk Tel: +44 (0)121 414 5439

Email: d.f.browning@bham.ac.uk Tel: +44 (0)121 414 5434

Supplementary Table S1. Bacterial strains and plasmid used in this study.

Name	Details	Source
Strains		
BL21	*fhuA2 [lon] ompT gal [dcm] ΔhsdS*	Novagen
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Novagen
JCB387	*Δnir, Δlac*	[1]
JWD3	*E. coli* K-12 BamA depletion strain	[2]
SHuffle Express	*F´ lac, pro, lacI^q / Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt::pNEB3-r1-cDsbC (Spec^R, lacI^q) ΔtrxB rpsL150(Str^R) Δgor Δ(malF) λ$^-$, IN(rrnD-rrnE)1, rph-1*	NEB
W3110	*recA1, endA1, gyrA96, thi-1, hsdR17, supE44,*	[3]
XL1 Blue	*relA1, lac, [F proAB lacI^{q}ZΔM15 Tn10 (Tet^R)]*	Agilent
Plasmids		
pET15b	T7 RNA polymerase expression vector (Amp^R, N-terminal His tag). Carries the *lacI* gene.	Novagen
pET22b	T7 RNA polymerase expression vector (Amp^R, *pelB* signal sequence, C-terminal His tag). Carries the *lacI* gene.	Novagen
pET26b	T7 RNA polymerase expression vector (Kan^R, *pelB* signal sequence, C-terminal His tag). Carries the *lacI* gene.	Novagen
pET15b/ *6his-gfp*	pET15b expressing 6His N-terminal GFP fusion	Dr David Lee.
pYU49	pET23 based vector with *ptac* promoter expressing TorAsp anti-IL-1β-6His scFv.	[4]
pHAK1	pET23 based vector with *ptac* promoter expressing TorAsp hGH-6His	[5]
pRW50	*lacZ* transcription fusion plasmid (Tet^R)	[6]
pRW50/ PAR1	pRW50 carrying the PAR1 promoter	[7]
pRW50/ PAR2	pRW50 carrying the PAR2 promoter	[7]
pRW50/ PAR3	pRW50 carrying the PAR3 promoter	[7]
pRW50/ PAR4	pRW50 carrying the PAR4 promoter	[7]
pRW50/ PAR4L	pRW50 carrying the PAR4L promoter	[7]
pRW50/ PAR5	pRW50 carrying the PAR5 promoter	[7]
pRW50/ PAR6	pRW50 carrying the PAR6 promoter	[7]
pRW50/ PAR7	pRW50 carrying the PAR7 promoter	[7]
pRW50/ PAR8	pRW50 carrying the PAR8 promoter	[7]

Supplementary Table S2. Primers used in this study

Primer	Sequence (5` to 3`)
ptac(BglII)	GGGGG**AGATCT**GATAATGTTTTTTGCGCCGACATCATAACGG
ptacO3(BglII)	GGGGG**AGATCT**GGCAGTGAGCGCAACGCAATTATCATAACGGTTCTGGC
placO1(BglII)	GGGGG**AGATCT**AATTGTGAGCGGATAACAATTAATGTGAGTTAGCTCACTC
placO3(BglII)	GGGGG**AGATCT**GGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTC
placRV(XbaI)	GGGGG**TCTAGA**CTGTTTCCTGTGTGAAATTGTTATCCG
lacIqF	GACACCATCGAATGG**T**GCAAAACCTTTCGCG
lacIqR	CGCGAAAGGTTTTGC**A**CCATTCGATGGTGTC

Supplementary Figure Legends.

Supplementary Fig. S1. The pET22b and pET26b multiple cloning site (MCS). The figure shows the DNA sequence of the T7 promoter region and MCS of pET22b and pET26b (Novagen). Note that the same region in each plasmid is identical. The location of the 6His purification tag and the *pelB* leader sequence, which allows secretion of protein into the *E. coli* periplasm, is highlighted. Restriction enzyme recognition sites are bold. The location of the T7 promoter, the *lac* operator, the ribosome binding site (RBS) and the T7 terminator primer (Novagen) are also indicated. Amino acid sequence is also shown below the relevant DNA sequence.

Supplementary Fig. S2. The DNA sequences of target proteins expressed in this study. The figure shows the DNA sequences of (a) *6his-gfp* (b) *hgh-6his* (c) *IL-1β-6his scFv* and (d) $bamA_{ENm}$ used in this study. Restriction enzyme recognition sites (XbaI, NdeI, BamHI, SacI and XhoI) used to clone each fragment into vectors are shown bold and underlined. For each construct the translation initiation codon (AUG) is green, the DNA encoding the 6His tag is purple and the translation stop codon (TAA) is red. Note the $bamA_{ENm}$ construct used in this work encodes the $BamA_{ENm}$ chimera protein, which is a fusion of the *E. coli* BamA N-terminal domain and BamA C-terminal β-barrel domain from *N. meningitidis* [8].

Supplementary Fig. S3. The amino acid sequences of target proteins expressed in this study. The figure shows the amino acid sequences of the (a) 6His-GFP (b) hGH-6His (c) anti-IL-1β-6His scFv and (d) $BamA_{ENm}$ proteins used in this study. For each protein the 6His tag is purple and predicted molecular weight (Mw) of each protein is given. Note the $BamA_{ENm}$ protein chimera is a fusion of the *E. coli* BamA N-terminal domain and BamA C-terminal β-barrel domain from *N. meningitidis* [8].

Supplementary Fig. S4. Expression of 6His-GFP using the PAR promoter system in the pET26b vector backbone. The figure shows a Coomassie blue stained SDS-PAGE gel analysing the expression of 6His-GFP from the three promoter PAR system (PAR1, PAR4L and PAR7Q) cloned into the pET26b vector system in *E. coli* BL21 cells. Cells were grown in LB medium at 37°C and RPP was induced for three hours by the addition of 1 mM IPTG, where appropriate. An empty vector control (EV) was included.

Supplementary Fig. S5. Expression of 6His-GFP using the PAR promoter system in *E. coli* K-12 strain W3110. The figure shows a Coomassie blue stained SDS-PAGE gel analysing the expression of 6His-GFP from the three promoter PAR system (PAR1, PAR4L and PAR7Q) cloned into the pET22b vector in *E. coli* K-12 strain W3110. Cells were grown in LB medium

at 37°C and RPP was induced for 3 hours by the addition of 1 mM IPTG, where appropriate. An empty vector control (EV) was included.

Supplementary Fig. S6. Expression of 6His-GFP using the PAR promoter system in auto-induction medium at different growth temperatures. The figure shows Coomassie blue stained SDS-PAGE gels analysing the expression of 6His-GFP in *E. coli* BL21 cells, using the three promoter PAR system (PAR1, PAR4L and PAR7Q) when cloned into pET22b. Cells were grown in auto-induction medium [9] at either (a) 37°C or (b) 30°C and samples were taken 3, 6 and 23 hours after sub-culturing (O/N overnight).

Supplementary Fig. S7. Homogeneous green fluorescence intensity of cells expressing 6His-GFP from the PAR promoters. Flow cytometry analysis of green fluorescence from BL21 pET22b PAR/ 6His-GFP constructs (a) PAR1, (b) PAR4L and (c) PAR7Q grown in LB medium with 2 µM to 1000 µM IPTG induction for four hours. Data are plotted as histograms showing number of cells with different green fluorescence (FL1-A) values.

Supplementary Fig. S8. Glucose represses expression from the PAR1 promoter. The figure shows the mean green fluorescence intensities of *E. coli* BL21 cells, carrying pET22b PAR1/ *6his-gfp*, for the first three hours after IPTG induction, in the presence and absence of 0.5% glucose. Dotted lines correspond to expression levels in cells grown in the absence of glucose, whilst the solid ones represent cells grown in its presence. 6His-GFP expression was induced by the inclusion of IPTG at 2, 10, 50, 100 and 1000 µM. Data are shown as mean green fluorescence values from replica flasks, error bars are ± the standard deviation.

Supplementary Fig. S9. Rescue of BamA depletion by the expression of the *N. meningitidis* $BamA_{ENm}$ chimera protein. (a) The panel shows growth of the *E. coli* BamA depletion strain, JWD3, on LB agar pates in the presence of arabinose (+Ara), the absence of arabinose (-Ara) or the presence of only 1 mM IPTG (+IPTG). Cells carried either pET22b PAR1 empty vector or pET22b PAR1/ $BamA_{ENm}$. Note that in JWD3 cells the chromosomally-encoded *E. coli* BamA is only produced in the presence of arabinose, whilst in its absence, BamA expression is prevented, resulting in cell death [2]. Depletion can be rescued by providing a functional copy of *bamA* [8, 10], such as that carried by pET22b PAR1/ $BamA_{ENm}$. (b) The panel shows the growth of JWD3 cells, carrying either pET22b PAR1 empty vector or pET22b PAR1/ $BamA_{ENm}$, in Lennox broth medium, supplemented with 0.2% glucose, in the absence of arabinose, the presence of arabinose (Ara) or the presence of 1 mM IPTG (IPTG). (c) Detection of $BamA_{ENm}$ chimera expression. The panel shows a Coomassie blue stained gel

672 and Western blot of normalised total cell protein from the JWD3 cells carrying either pET22b
673 PAR1 empty vector or pET22b PAR1 $BamA_{ENm}$, after 480 minutes of growth in Lennox broth
674 with 0.2% glucose supplemented with 0.05% arabinose or 1 mM IPTG, where indicated. Blots
675 were probed with anti-*E. coli* BamA POTRA antiserum and anti-rabbit HRP secondary
676 antibody to detect $BamA_{ENm}$.
677
678 **Supplementary Fig. S10.** Comparison of 6His-GFP production using standard T7 RNA
679 polymerase-driven expression and the PAR7Q promoter construct. The figure shows a
680 Coomassie blue stained SDS-PAGE gel analysing the expression of 6His-GFP in *E. coli*
681 BL21(DE3) cells, using pET15b/ *6his-gfp*, and in *E. coli* BL21, using pET22b PAR7Q/ *6his-*
682 *gfp*. Cells were grown in LB medium at 37°C with shaking and recombinant protein production
683 was induced for three hours by the addition of 1 mM IPTG, where appropriate.

Supplementary References

[1] Page L, Griffiths L, Cole JA. Different physiological roles of two independent pathways for nitrite reduction to ammonia by enteric bacteria. Archives of microbiology. 1990;154(4):349-54. PubMed PMID: 2173895; eng.

[2] Lehr U, Schütz M, Oberhettinger P, et al. C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. Molecular microbiology. 2010 Nov;78(4):932-46. doi: 10.1111/j.1365-2958.2010.07377.x. PubMed PMID: 20815824; eng.

[3] Hayashi K, Morooka N, Yamamoto Y, et al. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. Molecular systems biology. 2006;2:2006.0007. doi: 10.1038/msb4100049. PubMed PMID: 16738553; PubMed Central PMCID: PMCPmc1681481. Eng.

[4] Matos CF, Robinson C, Alanen HI, et al. Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in *Escherichia coli* CyDisCo strains. Biotechnology progress. 2014 Mar-Apr;30(2):281-90. doi: 10.1002/btpr.1858. PubMed PMID: 24376243; eng.

[5] Alanen HI, Walker KL, Lourdes Velez Suberbie M, et al. Efficient export of human growth hormone, interferon alpha2b and antibody fragments to the periplasm by the *Escherichia coli* Tat pathway in the absence of prior disulfide bond formation. Biochimica et biophysica acta. 2015 Mar;1853(3):756-63. doi: 10.1016/j.bbamcr.2014.12.027. PubMed PMID: 25554517; eng.

[6] Lodge J, Fear J, Busby S, et al. Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. FEMS microbiology letters. 1992 Aug 15;74(2-3):271-6. PubMed PMID: 1526459; eng.

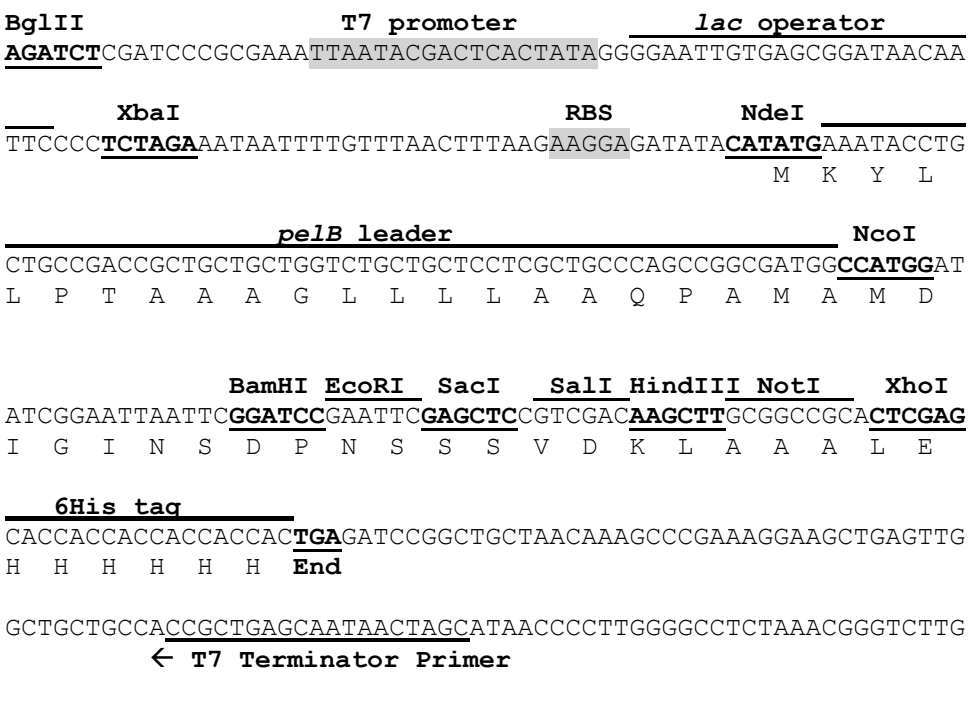
[7] Browning DF, Godfrey RE, Richards KL, et al. Exploitation of the *Escherichia coli lac* operon promoter for controlled recombinant protein production. Biochemical Society transactions. 2019 Apr 30;47(2):755-763. doi: 10.1042/bst20190059. PubMed PMID: 30971435; eng.

[8] Browning DF, Bavro VN, Mason JL, et al. Cross-species chimeras reveal BamA POTRA and beta-barrel domains must be fine-tuned for efficient OMP insertion. Molecular microbiology. 2015 Aug;97(4):646-59. doi: 10.1111/mmi.13052. PubMed PMID: 25943387; PubMed Central PMCID: PMCPmc4950039. Eng.

[9] Studier FW. Protein production by auto-induction in high density shaking cultures. Protein expression and purification. 2005 May;41(1):207-34. doi: 10.1016/j.pep.2005.01.016. PubMed PMID: 15915565; eng.

[10] Browning DF, Matthews SA, Rossiter AE, et al. Mutational and topological analysis of the *Escherichia coli* BamA protein. PloS one. 2013;8(12):e84512. doi: 10.1371/journal.pone.0084512. PubMed PMID: 24376817; PubMed Central PMCID: PMCPmc3871556. Eng.

Supplementary Fig. S1.



Supplementary Fig. S2.

(a) *6his-gfp*

TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACC**ATG**GGCAGCAGC**CATCATCATCATCATCAC**AGCA
GCGGCCTGGTGCCGCGCGGCAGC**CATATG**GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGG
TCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA
AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCT
ACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAG
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGG
GCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGC
TGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCA
AGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG
GCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCG
ATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG**TAA**GGAT
CC

(b) *hgh-6his*

CATATGTTCCCAACCATTCCCTTATCCAGGCTTTTTGACAACGCTATGCTCCGCGCCCATCGTCTGCACCAGCTG
GCCTTTGACACCTACCAGGAGTTTGAAGAAGCCTATATCCCAAAGGAACAGAAGTATTCATTCCTGCAGAACCCC
CAGACCTCCCTCTGTTTCTCAGAGTCTATTCCGACACCCTCCAACAGGGAGGAAACACAACAGAAATCCAACCTA
GAGCTGCTCCGCATCTCCCTGCTGCTCATCCAGTCGTGGCTGGAGCCCGTGCAGTTCCTCAGGAGTGTCTTCGCC
AACAGCCTGGTGTACGGCGCCTCTGACAGCAACGTCTATGACCTCCTAAAGGACCTAGAGGAAGGCATCCAAACG
CTGATGGGGAGGCTGGAAGATGGCAGCCCCCGGACTGGGCAGATCTTCAAGCAGACCTACAGCAAGTTCGACACA
AACTCACACAACGATGACGCACTACTCAAGAACTACGGGCTGCTCTACTGCTTCAGGAAGGACATGGACAAGGTC
GAGACATTCCTGCGCATCGTGCAGTGCCGCTCTGTGGAGGGCAGCTGTGGCTTC**CATCATCATCATCATCAC****TAA**
TAAGGATCCGAATTC**GAGCTC**

(c) *IL-1β-6his scFv*

CATATGGATATCCAGATGACGCAGAGCCCGAGCAGCCTGAGCGCCAGCGTGGGTGACCGTGTGACCATTACCTGT
CGTACCAGCGGCAACATTCATAACTATCTGACCTGGTACCAGCAGAAACCGGGCAAAGCGCCGCAGCTGCTGATT
TATAATGCAAAAACCCTGGCAGATGGTGTGCCGAGCCGCTTTAGCGGCAGCGGTAGCGGTACCCAGTTCACCCTG
ACGATCAGCAGCCTGCAGCCGGAAGACTTTGCCAACTATTACTGCCAGCACTTCTGGAGCCTGCCGTTTACCTTC
GGTCAGGGCACGAAAGTGGAAATTAAACGTACCGGCGGTGGCGGTAGCGGCGGTGGCGGTAGCGGCGGTGGCGGT
AGCGGCGGTGGCGGTAGCGAGGTGCAGCTGGTTGAAAGCGGCGGTGGCCTGGTTCAGCCGGGTGGCAGCCTGCGT
CTGAGCTGTGCGGCCAGCGGCTTTGATTTCAGCCGTTATGACATGAGCTGGGTGCGTCAGGCACCGGGTAAACGT
CTGGAATGGGTTGCCTATATTAGCAGCGGTGGCGGTAGCACCTACTTTCCGGATACGGTGAAAGGCCGCTTCACC
ATCAGCCGTGACAACGCAAAAAATACGCTGTACCTGCAGATGAACAGCCTGCGCGCCGAAGATACCGCAGTTTAT
TACTGCGCCCGTCAGAATAAAAAACTGACGTGGTTCGACTACTGGGGTCAGGGCACGCTGGTTACGGTTAGCAGC
CATCATCATCATCATCAC**TAATAA**GGATCCGAATTC**GAGCTC**

(d) *bamA*$_{ENm}$

CATATGGCGATGAAAAAACTGCTGATCGCGTCTCTGCTGTTCTCTTCTGCGACCGTTTACGGTGCTAGCGAAGGT
TTCGTTGTTAAAGACATCCACTTTGAAGGTCTGCAACGTGTTGCGGTTGGTGCGGCGCTGCTGTCTATGCCGGTT
CGTACCGGCGACACCGTGAACGACGAAGACATCTCTAACACCATCCGTGCGCTGTTCGCGACCGGCAACTTTGAA
GACGTTCGTGTTCTGCGTGACGGTGACACCCTGCTGGTTCAGGTTAAAGAACGTCCGACCATCGCGTCTATCACC
TTCTCTGGTAACAAATCTGTTAAAGACGACATGCTGAAACAGAACCTGGAAGCGTCTGGTGTTCGTGTTGGTGAA
TCTCTGGACCGTACCACCATCGCGGACATCGAAAAAGGTCTGGAAGACTTCTACTACTCTGTTGGTAAATACTCT
GCGTCTGTTAAAGCGGTTGTTACCCCGCTGCCGCGTAACCGTGTTGACCTGAAACTGGTTTTCCAGGAAGGTGTT
TCTGCGGAAATCCAGCAGATCAACATCGTTGGTAACCACGCTTTCACCACCGACGAACTGATCTCTCACTTCCAA
CTGCGTGACGAAGTTCCGTGGTGGAACGTGGTTGGTGACCGTAAATACCAGAAACAGAAACTGGCGGGTGACCTG

GAAACCCTGCGTTCTTACTACCTGGACCGTGGTTACGCGCGTTTCAACATCGACTCTACCCAGGTTTCTCTGACC
CCGGACAAAAAAGGTATCTACGTTACCGTGAACATCACCGAAGGTGACCAGTACAAACTGTCTGGTGTTGAAGTT
TCTGGTAACCTGGCGGGTCACTCTGCGGAAATCGAACAACTGACCAAAATCGAACCGGGTGAACTGTATAACGGC
ACCAAAGTTACCAAAATGGAAGACGACATCAAAAAACTGCTGGGTCGTTACGGTTACGCTTACCCGCGTGTTCAG
TCTATGCCGGAAATCAACGACGCGGACAAAACCGTTAAACTGCGTGTGAACGTGGACGCGGGTAACCGTTTCTAC
GTTCGTAAAATCCGTTTTGAAGGTAACGACACCTCTAAAGACGCGGTTCTGCGTCGTGAAATGCGTCAGATGGAA
GGTGCGTGGCTGGGTTCTGACCTGGTTGACCAGGGTAAAGAACGTCTGAACCGTCTGGGTTTCTTTGAAACCGTT
GACACCGACACCCAGCGTGTTCCGGGTTCCCCGGACCAGGTTGACGTTGTTTACAAAGTTAAAGAACGTAACACC
GGATCCCTGGACCTGTCTGCGGGTTGGGTTCAGGACACCGGCCTGGTTATGTCTGCGGGTGTTTCTCAGGACAAC
CTGTTCGGCACCGGCAAATCTGCGGCGCTGCGTGCGTCTCGTTCTAAAACCACCCTGAACGGTTCTCTGTCTTTC
ACCGACCCGTACTTCACCGCTGACGGTGTTTCTCTGGGTTACGACGTTTACGGTAAAGCGTTCGACCCGCGTAAA
GCGTCTACCTCTATCAAACAGTACAAAACCACCACCGCTGGTGCGGGTATCCGTATGTCTGTTCCGGTTACCGAA
TACGACCGTGTGAACTTCGGTCTGGTTGCGGAACACCTGACCGTGAACACCTACAACAAAGCGCCGAAACACTAC
GCGGACTTCATCAAAAAATACGGTAAAACCGACGGCACCGACGGTTCTTTCAAAGGTTGGCTGTATAAAGGCACC
GTTGGTTGGGGTCGTAACAAAACCGACTCTGCGCTGTGGCCGACCCGTGGTTACCTGACCGGCGTGAACGCGGAA
ATCGCGCTGCCGGGTTCTAAACTGCAATACTACTCTGCGACCCACAACCAGACCTGGTTCTTCCCGCTGTCTAAA
ACCTTCACCCTGATGCTGGGTGGTGAAGTTGGTATCGCGGGTGGTTACGGTCGTACCAAAGAAATCCCGTTCTTT
GAAAACTTCTACGGTGGTGGTCTGGGTTCTGTTCGTGGTTACGAATCTGGCACCCTGGGTCCGAAAGTTTACGAC
GAATACGGTGAAAAAATCTCTTACGGTGGTAACAAAAAAGCGAACGTGTCTGCGGAACTGCTGTTCCCGATGCCG
GGTGCGAAAGACGCGCGTACCGTTCGTCTGTCTCTGTTCGCGGACGCGGGTTCTGTTTGGGACGGTAAAACCTAC
GACGACAACTCTTCTTCTGCGACCGGCGGTCGTGTTCAGAACATCTACGGTGCGGGTAACACCCACAAATCTACC
TTCACCAACGAACTGCGTTACTCTGCGGGTGGTGCGGTTACCTGGCTGTCTCCGCTGGGGCCCATGAAATTCTCT
TACGCTTACCCGCTGAAAAAAAAACCGGAAGACGAAATCCAGCGTTTCCAGTTCCAACTGGGCACCACCTTC**TAA**
TGAGGGCCCATGAAGTTTAGCTATGCCTATCCATTAAAGAAGAAGCCAGAGGATGAGATTCAAAGATTTCAATTT
CAATTAGGTACTACTTTTGGCGGCAGATCT**CTCGAG**

Supplementary Fig. S3.

(a) 6His-GFP Mw 29105 Da

MGSS**HHHHHH**SSGLVPRGSHMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ
QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

(b) hGH-6His Mw 23083 Da

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREET
QQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRT
GQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF**HHHHHH**

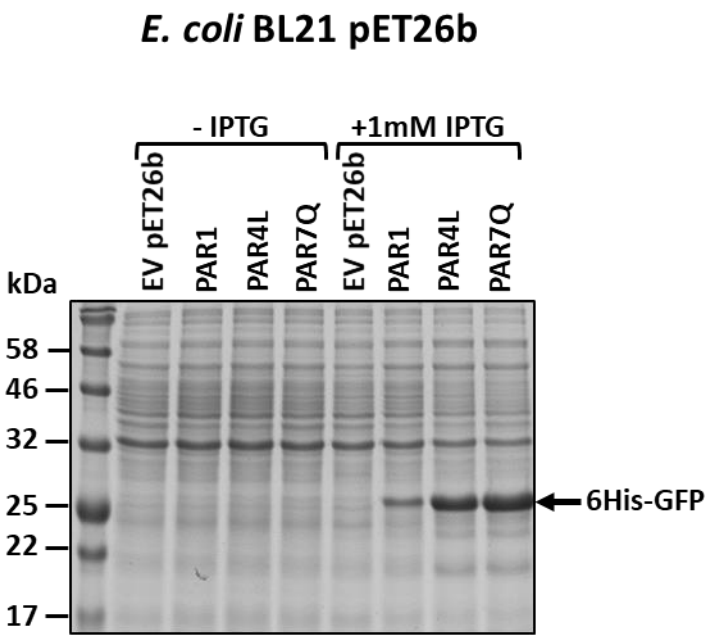
(c) Ant-IL-1β-6His scFv Mw 27495 Da

MDIQMTQSPSSLSASVGDRVTITCRTSGNIHNYLTWYQQKPGKAPQLLIYNAKTLADGVPSRFSGSGS
GTQFTLTISSLQPEDFANYYCQHFWSLPFTFGQGTKVEIKRTGGGGSGGGGSGGGGSGGGGSEVQLVE
SGGGLVQPGGSLRLSCAASGFDFSRYDMSWVRQAPGKRLEWVAYISSGGGSTYFPDTVKGRFTISRDN
AKNTLYLQMNSLRAEDTAVYYCARQNKKLTWFDYWGQGTLVTVSS**HHHHHH**

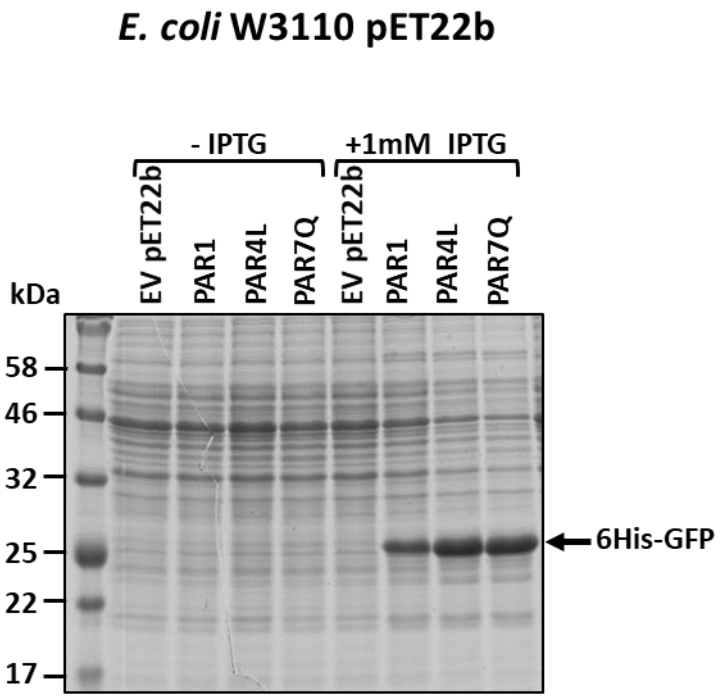
(d) BamA$_{ENm}$ Mw 87992 Da

MAMKKLLIASLLFSSATVYGASEGFVVKDIHFEGLQRVAVGAALLSMPVRTGDTVNDEDISNTIRALF
ATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVRVGESLDRTTIADIEK
GLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAEIQQINIVGNHAFTTDELISHFQLRDEV
PWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNIDSTQVSLTPDKKGIYVTVNITEGDQYKLSGV
EVSGNLAGHSAEIEQLTKIEPGELYNGTKVTKMEDDIKKLLGRYGYAYPRVQSMPEINDADKTVKLRV
NVDAGNRFYVRKIRFEGNDTSKDAVLRREMRQMEGAWLGSDLVDQGKERLNRLGFFETVDTDTQRVPG
SPDQVDVVYKVKERNTGSLDLSAGWVQDTGLVMSAGVSQDNLFGTGKSAALRASRSKTTLNGSLSFTD
PYFTADGVSLGYDVYGKAFDPRKASTSIKQYKTTTAGAGIRMSVPVTEYDRVNFGLVAEHLTVNTYNK
APKHYADFIKKYGKTDGTDGSFKGWLYKGTVGWGRNKTDSALWPTRGYLTGVNAEIALPGSKLQYYSA
THNQTWFFPLSKTFTLMLGGEVGIAGGYGRTKEIPFFENFYGGGLGSVRGYESGTLGPKVYDEYGEKI
SYGGNKKANVSAELLFPMPGAKDARTVRLSLFADAGSVWDGKTYDDNSSSATGGRVQNIYGAGNTHKS
TFTNELRYSAGGAVTWLSPLGPMKFSYAYPLKKKPEDEIQRFQFQLGTTF

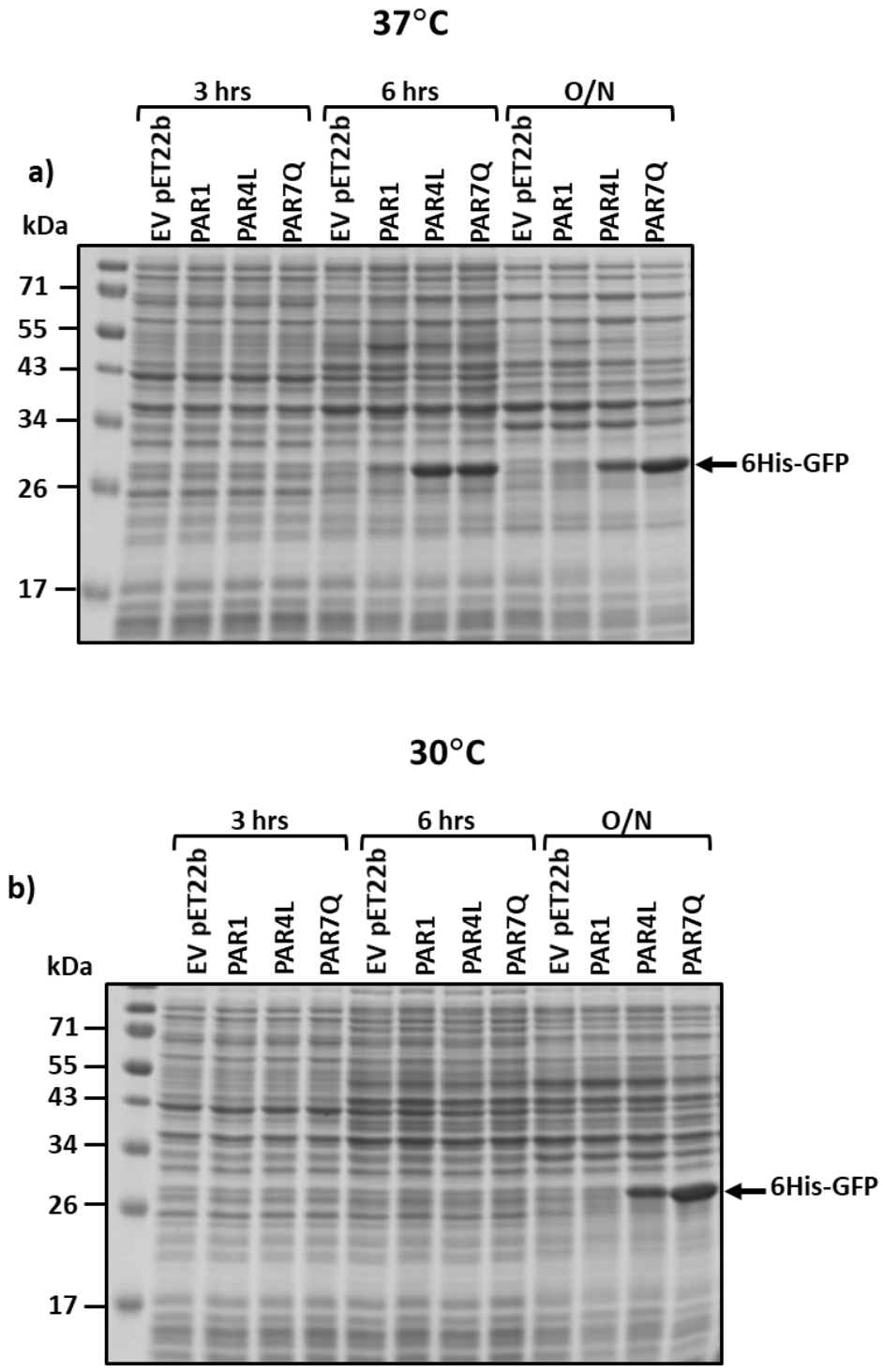
Supplementary Fig. S4.



Supplementary Fig. S5.



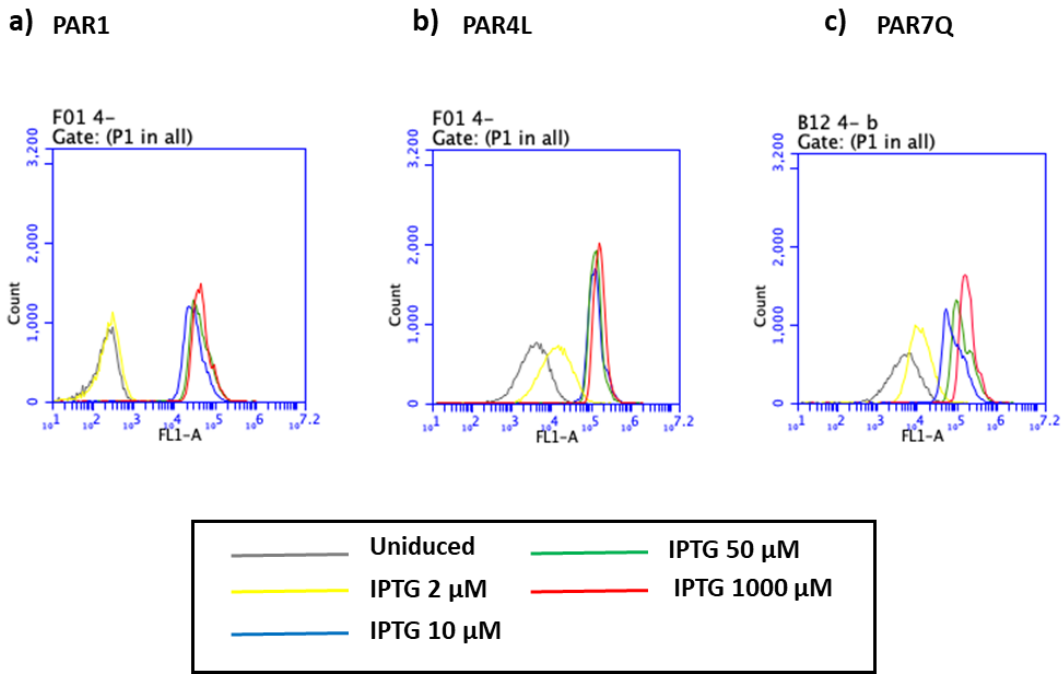
Supplementary Fig. S6.



878 **Supplementary Fig. S7.**

879

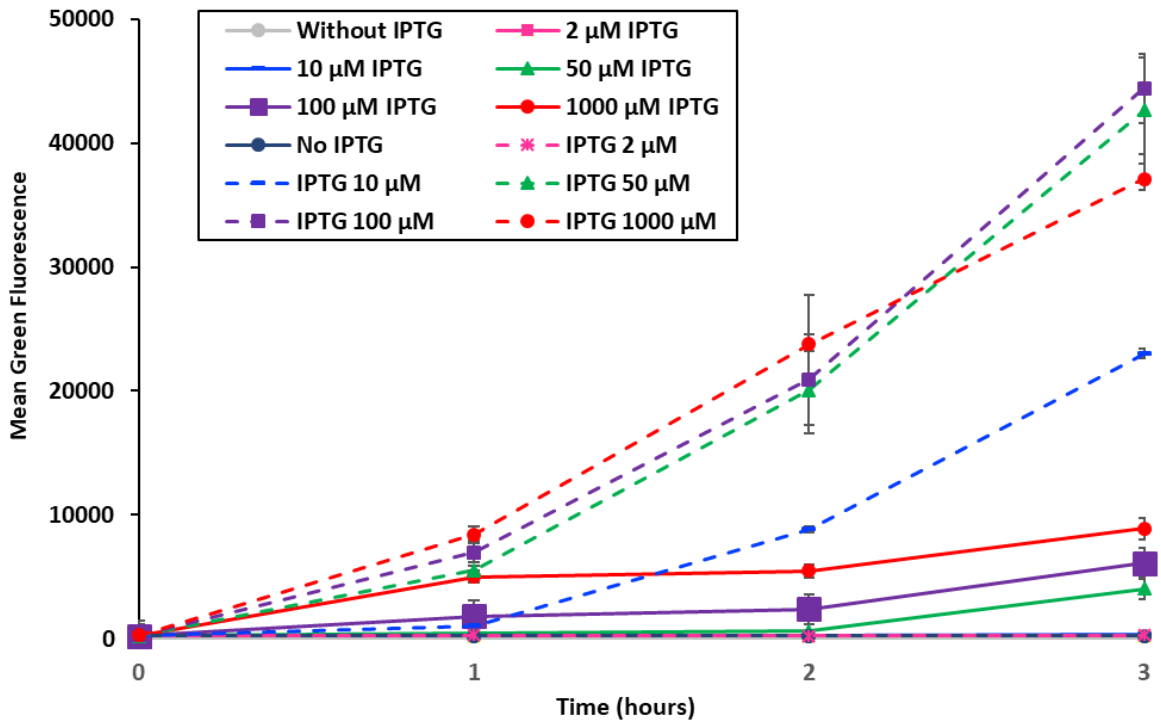
880



881

882 **Supplementary Fig. S8.**

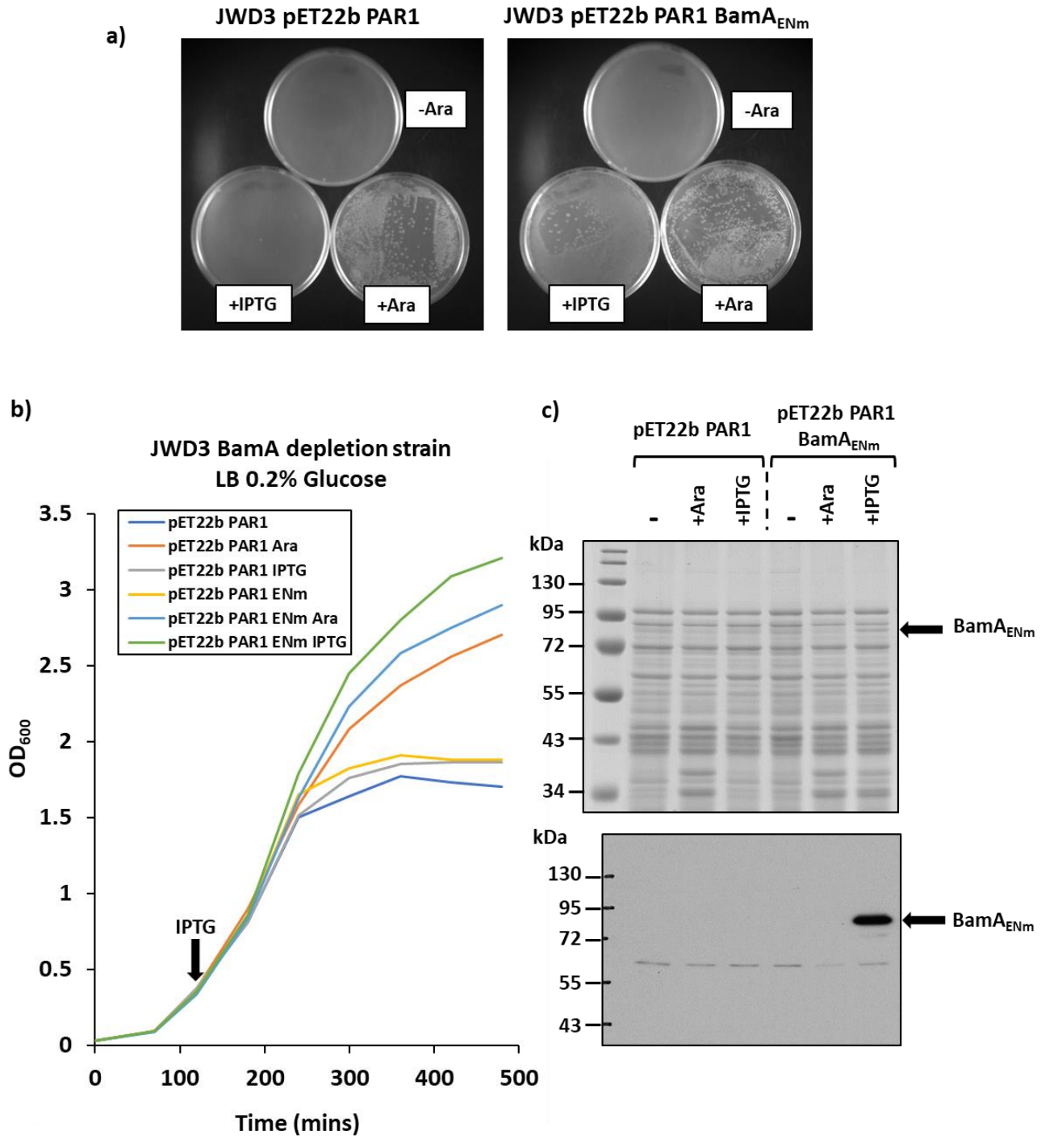
883



884

885

Supplementary Fig. S9.



Supplementary Fig. S10.